

1610
1657
OIPER 13
RAW SEQUENCE LISTING
PATENT APPLICATION US/09/492,697DATE: 02/16/2000
TIME: 02:14:57

INPUT SET: S34756.raw

This Raw Listing contains the General
Information Section and up to the first 5 pages.MAY 01 2000
TECH CENTER 1600/2901

SEQUENCE LISTING

ENTERED

1
2
3 (1) General Information:
4
5 (i) APPLICANT: Choulika, Andre
6 Perrin, Arnaud
7 Dujon, Bernard
8 Nicolas, Jean-Francois
9
10 (ii) TITLE OF INVENTION: Nucleotide Sequence Encoding the Enzyme
11 I-SCEI and the Uses Thereof
12
13 (iii) NUMBER OF SEQUENCES: 52
14
15 (iv) CORRESPONDENCE ADDRESS:
16 (A) ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
17 Dunner
18 (B) STREET: 1300 I Street, N.W.
19 (C) CITY: Washington
20 (D) STATE: D.C.
21 (E) COUNTRY: USA
22 (F) ZIP: 20005-3315
23
24 (v) COMPUTER READABLE FORM:
25 (A) MEDIUM TYPE: Floppy disk
26 (B) COMPUTER: IBM PC compatible
27 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
28 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
29
30 (vi) CURRENT APPLICATION DATA:
31 (A) APPLICATION NUMBER: Unassigned
32 (B) FILING DATE: 25-JAN-2000
33 (C) CLASSIFICATION:
34
35 (vii) PRIOR APPLICATION DATA:
36 (A) APPLICATION NUMBER: US 09/244,130
37 (B) FILING DATE: 04-FEB-1999
38
39 (vii) PRIOR APPLICATION DATA:
40 (A) APPLICATION NUMBER: US 09/119,024
41 (B) FILING DATE: 20-JUL-1998
42
43 (vii) PRIOR APPLICATION DATA:
44 (A) APPLICATION NUMBER: US 08/336,241
45 (B) FILING DATE: 07-NOV-1994
46

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/492,697

MAY 01 2000
TECH CENTER 1600/2900
DATE: 02/16/2000
TIME: 02:14:57

INPUT SET: S34756.raw

47 (vii) PRIOR APPLICATION DATA:
48 (A) APPLICATION NUMBER: US 07/971,160
49 (B) FILING DATE: 05-NOV-1992
50 (vii) PRIOR APPLICATION DATA:
51 (A) APPLICATION NUMBER: US 07/879,689
52 (B) FILING DATE: 05-MAY-1992
53 (viii) ATTORNEY/AGENT INFORMATION:
54 (A) NAME: Meyers, Kenneth J.
55 (B) REGISTRATION NUMBER: 25,146
56 (C) REFERENCE/DOCKET NUMBER: 3495-0111-11
57
58 (ix) TELECOMMUNICATION INFORMATION:
59 (A) TELEPHONE: 202-408-4000
60 (B) TELEFAX: 202-408-4400.
61
62
63
64 (2) INFORMATION FOR SEQ ID NO:1:
65
66 (i) SEQUENCE CHARACTERISTICS:
67 (A) LENGTH: 714 base pairs
68 (B) TYPE: nucleic acid
69 (C) STRANDEDNESS: single
70 (D) TOPOLOGY: linear
71
72 (ii) MOLECULE TYPE: DNA (genomic)
73
74
75
76 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
77
78 ATGCATATGA AAAACATCAA AAAAAACCAAG GTAATGAACC TCGGTCCGAA CTCTAACTG 60
79
80 CTGAAAGAAT ACAAAATCCC GCTGATCGAA CTGAACATCG AACAGTTCGA AGCAGGTATC 120
81
82 GGTCTGATCC TGGGTGATGC TTACATCCGT TCTCGTGATG AAGGTAAAC CTACTGTATG 180
83
84 CAGTCGAGT GGAAAAACAA AGCATAACATG GACCACGTAT GTCTGCTGTA CGATCAGTGG 240
85
86 GTACTGTCCC CGCCGCACAA AAAAGAACGT GTTAACCACC TGGTAAACCT GGTAATCACC 300
87
88 TGGGGCGCCC AGACTTCAA ACACCAAGCT TTCAACAAAC TGGCTAACCT GTTCATCGTT 360
89
90 AACAAACAAAA AAACCATCCC GAACAACCTG GTTGAAAAC ACCTGACCCC GATGTCTCTG 420
91
92 GCATACTGGT TCATGGATGA TGGTGGTAA TGTTGATTACA ACAAAAACTC TACCAACAAA 480
93
94 TCGATCGTAC TGAACACCCA GTCTTCACT TTCAAGAAC TAGAATACCT GTTTAAGGGT 540
95
96 CTGCGTAACA AATTCCAAC TGAACGTTAC GTAAAAATCA ACAAAAACAA ACCGATCATC 600
97
98 TACATCGATT CTATGTCTTA CCTGATCTTC TACAACCTGA TCAAACCGTA CCTGATCCG 660
99

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/492,697DATE: 02/16/2000
TIME: 02:14:57

INPUT SET: S34756.raw

100 CAGATGATGT ACAAACTGCC GAACACTATC TCCTCCGAAA CTTTCCTGAA ATAA 714

101

102 (2) INFORMATION FOR SEQ ID NO:2:

103

104 (i) SEQUENCE CHARACTERISTICS:

105 (A) LENGTH: 237 amino acids

106 (B) TYPE: amino acid

107 (D) TOPOLOGY: linear

108

109 (ii) MOLECULE TYPE: peptide

110

111

112

113 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

114

115 Met His Met Lys Asn Ile Lys Lys Asn Gln Val Met Asn Leu Gly Pro 15

116 1 5 10

117

118 Asn Ser Lys Leu Leu Lys Glu Tyr Lys Ser Gln Leu Ile Glu Leu Asn 30

119 20 25

120

121 Ile Glu Gln Phe Glu Ala Gly Ile Gly Leu Ile Leu Gly Asp Ala Tyr 45

122 35 40

123

124 Ile Arg Ser Arg Asp Glu Gly Lys Thr Tyr Cys Met Gln Phe Glu Trp 60

125 50 55

126

127 Lys Asn Lys Ala Tyr Met Asp His Val Cys Leu Leu Tyr Asp Gln Trp 80

128 65 70 75

129

130 Val Leu Ser Pro Pro His Lys Lys Glu Arg Val Asn His Leu Gly Asn 95

131 85 90

132

133 Leu Val Ile Thr Trp Gly Ala Gln Thr Phe Lys His Gln Ala Phe Asn 110

134 100 105

135

136 Lys Leu Ala Asn Leu Phe Ile Val Asn Asn Lys Thr Ile Pro Asn 125

137 115 120

138

139 Asn Leu Val Glu Asn Tyr Leu Thr Pro Met Ser Leu Ala Tyr Trp Phe 140

140 130 135

141

142 Met Asp Asp Gly Gly Lys Trp Asp Tyr Asn Lys Asn Ser Thr Asn Lys 160

143 150 155

144

145 Ser Ile Val Leu Asn Thr Gln Ser Phe Thr Phe Glu Glu Val Glu Tyr 175

146 165 170

147

148 Leu Val Lys Gly Leu Arg Asn Lys Phe Gln Leu Asn Cys Tyr Val Lys 190

149 180 185

150

151 Ile Asn Lys Asn Lys Pro Ile Ile Tyr Ile Asp Ser Met Ser Tyr Leu

152

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/492,697DATE: 02/16/2000
TIME: 02:14:58

INPUT SET: S34756.raw

153	195	200	205
154	Ile Phe Tyr Asn Leu Ile Lys Pro Tyr Leu Ile Pro Gln Met Met Tyr		
155	210	215	220
156			
157	Lys Leu Pro Asn Thr Ile Ser Ser Glu Thr Phe Leu Lys		
158	225	230	235
159			

160 (2) INFORMATION FOR SEQ ID NO:3:

162 (i) SEQUENCE CHARACTERISTICS:
 163 (A) LENGTH: 722 base pairs
 164 (B) TYPE: nucleic acid
 165 (C) STRANDEDNESS: single
 166 (D) TOPOLOGY: linear
 167
 168 (ii) MOLECULE TYPE: DNA (genomic)

169
 170
 171
 172 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

173	AAAAATAAAA TCATATGAAA AATATTAAAA AAAATCAAGT AATCAATCTC GGTCTTATT	60
174	CTAAATTATT AAAAGAATAT AAATCACAAT TAATTGAATT AAATATTGAA CAATTGAAAG	120
175	CAGGTATTGG TTTAATTGTT GGAGATGCTT ATATTCGTAG TCGTGATGAA GGTAAAAC	180
176	ATTGTATGCA ATTTGAGTGG AAAAATAAGG CATACTGGA TCATGTATGT TTATTATATG	240
177	ATCAATGGGT ATTATCACCT CCTCATAAAA AAGAAAGAGT TAATCATTAA GGTAATTAG	300
178	TAATTACCTG GGGAGCTCAA ACTTTAAC ATCAAGCTTT TAATAATTAA GCTAACTTAT	360
179	TTATTGTAAA TAATAAAAAA CTTATTCCTA ATAATTAGT TGAAAATTAT TTAACACCTA	420
180	TGAGTCTGGC ATATTGGTTT ATGGATGATG GAGGTAAATG GGATTATAAT AAAAATTCTC	480
181	TTAATAAAAG TATTGTATTA AATACACAAA GTTTACTTT TGAAGAAGTA GAATATTAC	540
182	TTAAAGGTTT AAGAAATAAA TTCAATTAA ATTGTATGT TAAAATTAAAT AAAAATAAAC	600
183	CAATTATTAA TATTGATTCT ATGAGTTATC TGATTTTTA TAATTAAATT AAACCTTATT	660
184	TAATTCCCTCA AATGATGTAT AACTGCCTA ATACTATTTC ATCCGAAACT TTTTAAAT	720
185	AA	722

199 (2) INFORMATION FOR SEQ ID NO:4:

200 (i) SEQUENCE CHARACTERISTICS:
 201 (A) LENGTH: 235 amino acids
 202 (B) TYPE: amino acid

203
 204
 205

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/492,697DATE: 02/16/2000
TIME: 02:14:58

INPUT SET: S34756.raw

206 (D) TOPOLOGY: linear
207
208 (ii) MOLECULE TYPE: peptide
209
210
211 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:
212
213 Met Lys Asn Ile Lys Lys Asn Gln Val Met Asn Leu Gly Pro Asn Ser
214 5 10 15
215 1
216 Lys Leu Leu Lys Glu Tyr Lys Ser Gln Leu Ile Glu Leu Asn Ile Glu
217 20 25 30
218
219 Gln Phe Glu Ala Gly Ile Gly Leu Ile Leu Gly Asp Ala Tyr Ile Arg
220 35 40 45
221
222 Ser Arg Asp Glu Gly Lys Thr Tyr Cys Met Gln Phe Glu Trp Lys Asn
223 50 55 60
224
225 Lys Ala Tyr Met Asp His Val Cys Leu Leu Tyr Asp Gln Trp Val Leu
226 65 70 80
227
228 Ser Pro Pro His Lys Lys Glu Arg Val Asn His Leu Gly Asn Leu Val
229 85 90 95
230
231 Ile Thr Trp Gly Ala Gln Thr Phe Lys His Gln Ala Phe Asn Lys Leu
232 100 105 110
233
234 Ala Asn Leu Phe Ile Val Asn Asn Lys Leu Ile Pro Asn Asn Leu
235 115 120 125
236
237 Val Glu Asn Tyr Leu Thr Pro Met Ser Leu Ala Tyr Trp Phe Met Asp
238 130 135 140
239
240 Asp Gly Gly Lys Trp Asp Tyr Asn Lys Asn Ser Leu Asn Lys Ser Ile
241 145 150 155 160
242
243 Val Leu Asn Thr Gln Ser Phe Thr Phe Glu Glu Val Cys Tyr Leu Val
244 165 170 175
245
246 Lys Gly Leu Arg Asn Lys Phe Gln Leu Asn Cys Tyr Val Lys Ile Asn
247 180 185 190
248
249 Lys Asn Lys Pro Ile Ile Tyr Ile Asp Ser Met Ser Tyr Leu Ile Phe
250

PAGE: 1

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION US/09/492,697

DATE: 02/16/2000
TIME: 02:14:58

INPUT SET: S34756.raw

Line	Error	Original Text
31	Wrong application Serial Number	(A) APPLICATION NUMBER: Unassigned